Notes on Sufficient Conditions for Extrema

Francis J. Narcowich
Department of Mathematics
Texas A&M University

January 2006

1 The second variation

Let $J[x] = \int_a^b F(t, x, \dot{x})dt$ be a nonlinear functional, with $x(a) = A$ and $x(b) = B$ fixed. As usual, we will assume that F is as smooth as necessary. The first variation of J is

$$\delta J_x[h] = \int_a^b \left(F(t, x, \dot{x}) - \frac{d}{dt} F_{\dot{x}} \right) h(t)dt,$$

where $h(t)$ is assumed as smooth as necessary and in addition satisfies $h(a) = h(b) = 0$. We will call such h *admissible*.

The idea behind finding the first variation is to capture the linear part of the $J[x]$. Specifically, we have

$$J[x + \varepsilon h] = J[x] + \varepsilon \delta J_x[h] + o(\varepsilon),$$

where $o(\varepsilon)$ is a quantity that satisfies

$$\lim_{\varepsilon \to 0} \frac{o(\varepsilon)}{\varepsilon} = 0.$$

The second variation comes out of the quadratic approximation in ε,

$$J[x + \varepsilon h] = J[x] + \varepsilon \delta J_x[h] + \frac{1}{2} \varepsilon^2 \delta^2 J_x[h] + o(\varepsilon^2).$$

It follows that

$$\delta^2 J_x[h] = \frac{d^2}{d\varepsilon^2} \Big(J[x + \varepsilon h] \Big) \Big|_{\varepsilon = 0}.$$

To calculate it, we note that

$$\frac{d^2}{d\varepsilon^2}\Big(J[x+\varepsilon h]\Big) = \int_a^b \frac{\partial^2}{\partial\varepsilon^2}\Big(F(t, x+\varepsilon h, \dot{x}+\varepsilon\dot{h})\Big)dt.$$

Applying the chain rule to the integrand, we see that

$$\begin{aligned}\frac{\partial^2}{\partial\varepsilon^2}\big(F(t, x+\varepsilon h, \dot{x}+\varepsilon\dot{h})\big) &= \frac{\partial}{\partial\varepsilon}\big(F_x h + F_{\dot{x}}\dot{h}\big) \\ &= F_{xx}h^2 + 2F_{x\dot{x}}h\dot{h} + F_{\dot{x}\dot{x}}\dot{h}^2,\end{aligned}$$

where the various derivatives of F are evaluated at $(t, x+\varepsilon h, \dot{x}+\varepsilon\dot{h})$. Setting $\varepsilon = 0$ and inserting the result in our earlier expression for the second variation, we obtain

$$\delta^2 J_x[h] = \int_a^b F_{xx}h^2 + 2F_{x\dot{x}}h\dot{h} + F_{\dot{x}\dot{x}}\dot{h}^2 dt\,.$$

Note that the middle term can be written as $2F_{x\dot{x}}h\dot{h} = F_{x\dot{x}}\frac{d}{dt}h^2$. Using this in the equation above, integrating by parts, and employing $h(a) = h(b) = 0$, we arrive at

$$\begin{aligned}\delta^2 J_x[h] &= \int_a^b \Big\{\big(\underbrace{F_{xx} - \frac{d}{dt}F_{x\dot{x}}}_{Q}\big)h^2 + \underbrace{F_{\dot{x}\dot{x}}}_{P}\dot{h}^2\Big\}dt \\ &= \int_a^b (P\dot{h}^2 + Qh^2)dt. \qquad (1)\end{aligned}$$

2 Legendre's trick

Ultimately, we are interested in whether a given extremal for J is a weak (relative) minimum or maximum. In the sequel we will always assume that the function $x(t)$ that we are working with is an extremal, so that $x(t)$ satisfies the Euler-Lagrange equation, $\frac{d}{dt}F_{\dot{x}} = F_x$, makes the first variation $\delta J_x[h] = 0$ for all h, and fixes the functions $P = F_{\dot{x}\dot{x}}$ and $Q = F_{xx} - \frac{d}{dt}F_{x\dot{x}}$. To be definite, we will always assume we are looking for conditions for the extremum to be a weak *minimum*. The case of a maximum is similar.

Let's look at the integrand $P\dot{h}^2 + Qh^2$ in (1). It is generally true that a function can be bounded, but have a derivative that varies wildly. Our intuition then says that $P\dot{h}^2$ is the dominant term, and this turns out to be true. In looking for a minimum, we recall that it is necessary that $\delta^2 J_x[h] \geq 0$

for all h. One can use this to show that, for a minimum, it is also necessary, but not sufficient, that $P \geq 0$ on $[a, b]$. We will make the stronger assumption that $P > 0$ on $[a, b]$. We also assume that P and Q are smooth.

Legendre had the idea to add a term to $\delta^2 J$ to make it nonnegative. Specifically, he added $\frac{d}{dt}(wh^2)$ to the integrand in (1). Note that $\int_a^b \frac{d}{dt}(wh^2)dt = wh^2|_a^b = 0$. Hence, we have this chain of equations,

$$\begin{aligned}
\delta^2 J_x[h] &= \delta^2 J_x[h] + \int_a^b \frac{d}{dt}(wh^2)dt \\
&= \int_a^b (P\dot{h}^2 + Qh^2 + \frac{d}{dt}(wh^2))dt \\
&= \int_a^b \left(P\dot{h}^2 + 2wh\dot{h} + (\dot{w} + Q)h^2\right) dt \qquad (2) \\
&= \int_a^b P\left(\dot{h} + \frac{w}{P}h\right)^2 dt + \int_a^b \left(\dot{w} + Q - \frac{w^2}{P}\right) h^2, \qquad (3)
\end{aligned}$$

where we completed the square to get the last equation. If we can find $w(t)$ such that

$$\dot{w} + Q - \frac{w^2}{P} = 0, \qquad (4)$$

then the second variation becomes

$$\delta^2 J_x[h] = \int_a^b P\left(\dot{h} + \frac{w}{P}h\right)^2 dt. \qquad (5)$$

Equation (4) is called a *Riccati* equation. It can be turned into the second order *linear* ODE below via the substitution $w = -(\dot{u}/u)P$:

$$-\frac{d}{dt}\left(P\frac{du}{dt}\right) + Qu = 0, \qquad (6)$$

which is called the *Jacobi* equation for J. Two points $t = \alpha$ and $t = \tilde{\alpha}$, $\alpha \neq \tilde{\alpha}$, are said to be *conjugate points* for Jacobi's equation if there is a solution u to (6) such that $u \not\equiv 0$ between α and $\tilde{\alpha}$, and such that $u(\alpha) = u(\tilde{\alpha}) = 0$.

When there are no points conjugate to $t = a$ in the interval $[a, b]$, we can construct a solution to (6) that is strictly positive on $[a, b]$. Start with the two linearly indepemdent solutions u_0 and u_1 to (6) that satsify the initial conditions

$$u_0(a) = 0,\ \dot{u}_0(a) = 1,\ u_1(a) = 0,\ \text{and}\ \dot{u}_1(a) = 1.$$

Since there is no point in $[a,b]$ conjugate a, $u_0(t) \neq 0$ for any $a < t \leq b$. In particular, since $\dot{u}_0(a) = 1 > 0$, $u(t)$ will be strictly positive on $(a,b]$. Next, because $u_1(a) = 1$, there exists $t = c$, $a < c \leq b$, such that $u_1(t) \geq 1/2$ on $[a,c]$. Moreover, the continuity of u_0 and u_1 on $[c,b]$ implies that $\min_{c \leq t \leq b} u_0(t) = m_0 > 0$ and $\min_{c \leq t \leq b} u_1(t) = m_1 \in \mathbb{R}$. It is easy to check that on $[a,b]$,

$$u := \frac{1+2|m_1|}{2m_0} u_0 + u_1 \geq 1/2,$$

and, of course, u solves (6).

This means that the substitutuion $w = -(\dot{u}/u)P$ yields a solution to the Riccati equation (4), and so the second variation has the form given in (5). It follows that $\delta^2 J_x[h] \geq 0$ for any admissible h. Can the second variation vanish for some h that is nonzero? That is, can we find an admissible $h \not\equiv 0$ such that $\delta^2 J_x[h] = 0$? If it did vanish, we would have to have

$$P\left(\dot{h} + \frac{w}{P}h\right)^2 = 0, \ a \leq t \leq b,$$

and, since $P > 0$, this implies that $\dot{h} + \frac{w}{P}h = 0$. This first order linear equation has the unique solution,

$$h(t) = h(a) e^{-\int_a^t \frac{w(\tau)}{P(\tau)} d\tau}.$$

However, since h is admissible, $h(a) = h(b) = 0$, and so $h(t) \equiv 0$. We have proved the following result.

Proposition 2.1. *If there are no points in $[a,b]$ conjugate to $t = a$, the the second variation is a positive definite quadratic functional. That is, $\delta^2 J_x[h] > 0$ for any admissible h not identically 0.*

3 Conjugate points

There is direct connection between conjugate points and extremals. Let $x(t,\varepsilon)$ be a family of extremals for the functional J depending smoothly on a parameter ε. We will assume that $x(a,\varepsilon) = A$, which will be independent of ε. These extremals all satisfy the Euler-Lagrange equation

$$F_x(t, x(t,\varepsilon), \dot{x}(t,\varepsilon) = \frac{d}{dt} F_{\dot{x}}(t, x(t,\varepsilon), \dot{x}(t,\varepsilon).$$

If we differentiate this equation with respect to ε, being careful to correctly apply the chain rule, we obtain

$$\begin{aligned} F_{xx}\frac{\partial x}{\partial \varepsilon} + F_{x\dot{x}}\frac{\partial \dot{x}}{\partial \varepsilon} &= \frac{d}{dt}\left(F_{x\dot{x}}\frac{\partial x}{\partial \varepsilon} + F_{\dot{x}\dot{x}}\frac{\partial \dot{x}}{\partial \varepsilon}\right) \\ &= \frac{dF_{x\dot{x}}}{dt}\frac{\partial x}{\partial \varepsilon} + F_{x\dot{x}}\frac{\partial \dot{x}}{\partial \varepsilon} + \frac{d}{dt}\left(F_{\dot{x}\dot{x}}\frac{\partial \dot{x}}{\partial \varepsilon}\right). \end{aligned}$$

Cancelling and rearranging terms, we obtain

$$\left(F_{xx} - \frac{d}{dt}F_{x\dot{x}}\right)\frac{\partial x}{\partial \varepsilon} - \frac{d}{dt}\left(F_{\dot{x}\dot{x}}\frac{\partial \dot{x}}{\partial \varepsilon}\right) = 0. \tag{7}$$

Set $\varepsilon = 0$ and let $u(t) = \frac{\partial x}{\partial \varepsilon}(t, 0)$. Observe that the functions in the equation above, which is called the variational equation, are just $P = F_{\dot{x}\dot{x}}$ and $Q = F_{xx} - \frac{d}{dt}F_{x\dot{x}}$. Consequently, (7) is simply the Jacobi equation (6). The difference here is that we always have the initial conditions,

$$\begin{cases} u(a) &= \dfrac{\partial x}{\partial \varepsilon}(a, 0) = \dfrac{\partial A}{\partial \varepsilon} = 0, \\ \dot{u}(a) &= \dfrac{\partial \dot{x}}{\partial \varepsilon}(a, 0) \neq 0. \end{cases}$$

We remark that if $\dot{u}(a) = 0$, then $u(t) \equiv 0$.

What do conjugate points mean in this context? Suppose that $t = \tilde{a}$ is conjugate to $t = a$. Then we have

$$\frac{\partial x}{\partial \varepsilon}(\tilde{a}, 0) = u(\tilde{a}) = 0,$$

which holds *independently* of how our smooth family of extremals was constructed. It follows that at $t = \tilde{a}$, we have $x(\tilde{a}, \varepsilon) = x(\tilde{a}, 0) + o(\varepsilon)$. Thus, the family either crosses again at $\tilde{a}$, or comes close to it, accumulating to order higher than ε there.

4 Sufficient conditions

A sufficient condition for an extremal to be a relative minimum is that the second variation be strongly positive definite. This means that there is a $c > 0$, which is independent of h, such that for all admissible h one has

$$\delta^2 J_x[h] \geq c\|h\|^2_{H^1},$$

where $H^1 = H^1[a,b]$ denotes the usual Sobolev space of functions with distributional derivatives in $L^2[a,b]$.

Let us return to equation (2), where we added in terms depending on an arbitrary function w. In the integrand there, we will add and subtract $\sigma P\dot{h}^2$, where σ is an arbitary constant. The only requirement for now is that $0 < \sigma < \min_{t\in[a,b]} P(t)$. The result is

$$\delta^2 J_x[h] = \int_a^b \left((P-\sigma)\dot{h}^2 + 2wh\dot{h} + (\dot{w}+Q)h^2\right) dt + \sigma \int_a^b \dot{h}^2 dt\,.$$

For the first integral in the term on the right above, we repeat the argument that was used to arrive at (5). Everything is the same, except that P is replaced by $P-\sigma$. We arrive at this:

$$\begin{aligned} \delta^2 J_x[h] = & \int_a^b (P-\sigma)\left(\dot{h} + \frac{w}{P-\sigma}h\right)^2 dt \\ & + \int_a^b \left(\dot{w} + Q - \frac{w^2}{P-\sigma}\right) h^2 + \sigma \int_a^b \dot{h}^2 dt\,. \end{aligned} \tag{8}$$

We continue as we did in section 2. In the end, we arrive at the new Jacobi equation,

$$-\frac{d}{dt}\left((P-\sigma)\frac{du}{dt}\right) + Qu = 0\,. \tag{9}$$

The point is that if for the Jacobi equation (6) there are no points in $[a,b]$ conjugate to a, then, because the solutions are continuous functions of the parameter σ, we may choose σ small enough so that for (9) there will be no points conjugate to a in $[a,b]$. Once we have fouund σ small enough for this to be true, we fix it. We then solve the corresponding Riccati equation and employ it in (8) to obtain

$$\begin{aligned} \delta^2 J_x[h] &= \int_a^b (P-\sigma)\left(\dot{h} + \frac{w}{P-\sigma}h\right)^2 dt + \sigma \int_a^b \dot{h}^2 dt \\ &\geq \sigma \int_a^b \dot{h}^2 dt\,. \end{aligned}$$

Now, for an admissble h, it is easy to show that $\int_a^b h^2 dt \leq \frac{(b-a)^2}{2}\int_a^b \dot{h}^2 dt$, so that we have

$$\|h\|_{H^1}^2 = \int_a^b h^2 dt + \int_a^b \dot{h}^2 dt \leq \left(1 + \frac{(b-a)^2}{2}\right)\int_a^b \dot{h}^2 dt.$$

Consequently, we obtain this inequality:

$$\delta^2 J_x[h] \geq \frac{\sigma}{1+\frac{(b-a)^2}{2}} \|h\|_{H^1}^2 = c\|h\|_{H^1}^2 \, ,$$

which is what we needed for a relative minimum. We summarize what we found below.

Theorem 4.1. *A sufficient condition for an extremal $x(t)$ to be a relative minimum for the functional $J[x] = \int_a^b F(t, x, \dot{x})dt$, where $x(a) = A$ and $x(b) = B$, is that $P(t) = F_{\dot{x}\dot{x}}(t, x, \dot{x}) > 0$ for $t \in [a, b]$ and that the interval $[a, b]$ contain no points conjugate to $t = a$.*